

Weekly Report

1. 工作概述

- (1) 本周我测试了三个数据集在 learning to rank 上的表现。这三个数据集是给定一个中心点和一个查询集，找到查询集中距离中心点最近的点，标签是查询集中距离中心点最近的点。我们采用 pair-wise 的方法训练 Rank 网络，网络输出两个点中哪个点距离中心点最近。根据训练时所用到的每个查询中的数据的不同，我们把这三个数据集分为：a) 对于每个查询集，只用距离中心点最近的点和分别和每个本查询集的其它点构成的查询对，训练 Rank 网络, b) 只用每一查询集中距离中心点最近的点和本查询集中另外一点作为训练点, c) 用本查询集中任意两点作为训练集。实际中，获取 a、b 两类数据比较容易，获取 c 类数据极难，实际中发现，使用 a、b 两类数据的训练结果差距不大。
- (2) 构造地图匹配训练的 ground truth(已完成)。结果如图 1 所示，其中绿线代表路网，红线代表轨迹，黑线代表轨迹实际行驶过程中的道路，中间的连线代表从轨迹到道路的匹配。

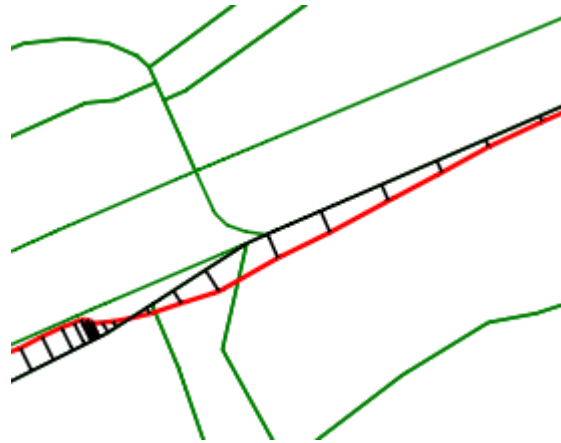


图 1 Ground Truth

- (3) 构造地图匹配的数据集(正在进行)。目前我们打算用轨迹某点 p_i 和其连续的 5 个点及道路的起始点作为网络的输入，输出为网络对结果的打分。我使用连续 5 个点作为训练数据是因为我们容易由连续的三个点计算出中点的行驶速度(假设匀速运动)，5 个点则可计算加速度，道路的起止点可用于计算道路的方向，距离的远近。最终输入数据的格式仍需要进一步优化。
- (4) 用有监督学习来解决地图匹配问题主要包括三个方面的过程：a) 选择、构建学习网络，使其计算模型能够通过学习方法来优化参数来逼近目标函数, b) 构造人工数据，检验自己对网络模型的结构、优化方面的判断无误, c) 将问题数据变换为网络的学习数据格式。其中耗时最大的地方是数据格式的变换。

2. 下周计划

- (1) 完成匹配结果
- (2) 完成论文初稿。可暂时先和局部匹配算法做对比

3. 论文阅读

暂无

4. 时间安排

Date	Tasks	Duration
Mon.	Reading and Programming	9:00-23:00

Work Time: above 50 hours